BAYESIAN TRACKING FOR FLUORESCENCE MICROSCOPIC IMAGING

Ihor Smal, Wiro Niessen and Erik Meijering

Biomedical Imaging Group Rotterdam
Erasmus MC – University Medical Center Rotterdam
Email: i.smal@erasmusmc.nl

ABSTRACT

Fluorescence microscopy is a powerful imaging tool for studying molecular dynamics in living cells. For quantitative motion analysis of subcellular structures robust and accurate detection and tracking techniques are necessary. Sequential Monte Carlo methods, also known as Particle Filters (PF), have become a tremendously popular tool to perform tracking in many fields. We propose a PF-based approach for quantitative analysis of subcellular dynamics. This approach utilizes all spatiotemporal information, which is an important advantage over existing methods that separate the object detection and object linking stage. The tracking technique has been evaluated using simulated but highly realistic image sequences, for which ground truth was available, showing that the method is more robust to noise than existing tracking techniques. In addition, evaluation experiments were conducted with real fluorescence microscopy image data acquired for specific biological studies.

1. INTRODUCTION

Advances in fluorescent probing and microscopic imaging technology have revolutionized biology in the past decade and have opened the door for studying subcellular dynamical processes. A simple description of sometimes complex patterns of movement in living cells may give insight in the underlying mechanisms governing these movements. However, accurate and reproducible methods for processing and analyzing the images acquired for such studies are still lacking. Since manual image analysis is time consuming, potentially inaccurate, and poorly reproducible, many biologically highly relevant questions are either left unaddressed, or answered with great uncertainty. Hence, the development of automated image analysis techniques for accurate and reproducible tracking and motion analysis of subcellular structures from time-lapse fluorescence microscopy image data is crucial.

Existing tracking techniques, whether commercial or academic, generally make very limited use of temporal information and prior knowledge. The conventional approach to tracking consists of two subsequent deterministic stages: detection (localization of objects) and linking (establishing correspondence between detected objects in adjacent frames). Detection methods include feature-based segmentation and model-based fitting, which do not always perform well on noisy microscopic imaging data [1]. Linking is based on the assumption that motion in natural images is mostly temporally smooth, that is, objects rarely change direction or speed abruptly, which is not always valid [2]. Tracking techniques based on optic flow [3] are not acceptable, because the underlying assumption that the total image intensity is constant over time is not satisfied in fluorescence microscopy due to photobleaching effects.

Recent results in psychophysics and human vision research have revealed the highly integrated nature of vision systems in using spatial, temporal, and prior information [4]. Local motion signals are often ambiguous, and many important motion phenomena can be explained by hypothesizing that the human visual system uses temporal coherence to resolve ambiguous inputs. It has therefore been proposed that input data are temporally grouped and used to predict and estimate the motion flows in image sequences. Temporal grouping can be expressed in terms of a Bayesian generalization of standard Kalman filtering.

In this paper the Bayesian tracking approach is implemented by a sequential Monte Carlo (MC) method. We demonstrate the applicability of the method for robust and accurate detection and tracking of large numbers of small objects in 2D and 3D image sequences obtained by fluorescence microscopy imaging.

2. METHODS

2.1. Bayesian Tracking

Bayesian estimation is used to recursively estimate a time evolving posterior distribution (also called filtering distribution) that describes the object state given all observations so far. It requires the definition of two models. The first one is a dynamic model that describes the evolution of the state $\{\mathbf{x}_k, k \in \mathbb{N}\}$ of the object given by

$$\mathbf{x}_k = \mathbf{f}_k\left(\mathbf{x}_{k-1}, \xi_{k-1}\right) \tag{1}$$

where $\mathbf{f}_k : \Re^{n_x} \times \Re^{n_\xi} \rightarrow \Re^{n_x}$ is a nonlinear function of the state $\mathbf{x}_{k-1}$ and $\{\xi_k, k \in \mathbb{N}\}$ is an independent and identically distributed (iid) process noise sequence. The n_x-dimensional state vector $\mathbf{x}_k$ may include the position, velocity and other parameters of the object. The second model is an observation model, which maps the true state space into the observed space and is given by

$$\mathbf{z}_k = \mathbf{h}_k\left(\mathbf{x}_k, \eta_k\right) \tag{2}$$

where $\mathbf{h}_k : \Re^{n_x} \times \Re^{n_\eta} \rightarrow \Re^{n_z}$ is a nonlinear function and $\{\eta_k, k \in \mathbb{N}\}$ is an iid measurement noise sequence. The observation $\mathbf{z}_k$ may include the image intensity and other features. This, in theory, is sufficient to allow recursive estimation of the filtering distribution, that is the calculation of some degree of belief in the state $\mathbf{x}_k$ at time k given the data $\{\mathbf{z}_j, j \in [1, k]\} \equiv \mathbf{z}_{1:k}$. The required probability density function (pdf) $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ may be obtained, recursively, in two stages: prediction and update. It is assumed that the initial pdf $p(\mathbf{x}_0|\mathbf{z}_0) \equiv p(\mathbf{x}_0)$, also known as the prior, is available ($\mathbf{z}_{1:0} = \mathbf{z}_0$ being the set of no measurements).

The prediction stage involves using the system model (1) and pdf $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$ to obtain the prior pdf of the state at time k via the Chapman-Kolmogorov equation

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1} \tag{3}$$

where the probabilistic model of the state evolution $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ is assumed to be Markovian and defined by the system equation (1) and the known statistics of ξ_k. In the update stage, when a measurement $\mathbf{z}_k$ becomes available, Bayes' rule is used in order to modify the prior density and obtain the required posterior density of the current state

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \propto p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) \tag{4}$$

where the likelihood function $p(\mathbf{z}_k|\mathbf{x}_k)$ is defined by the measurement model (2) and the known statistics η_k.

The recurrence relations (3) and (4) form the basis for the optimal Bayesian solution. This recursive propagation of the posterior density is only a conceptual solution because the likelihood models for tracking often lead to intractable inference, requiring approximation techniques. Solutions do exist in a restrictive set of cases, including the Kalman filter (models (1) and (2) are linear and Gaussian) and grid based filters.

2.2. Particle Filtering Methods

The practical implementation of the Bayesian estimation is usually achieved by Monte Carlo (MC) simulations and known variously as bootstrap filtering, particle filtering and the condensation algorithm [5]. It represents the required posterior density function $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ with a set of N_s random samples, or particles, and associated weights $\{\mathbf{x}_k^i, w_k^i\}_{i=1}^{N_s}$. Thus, the filtering distribution can be approximated as

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) \approx \sum_{i=1}^{N_s} w_k^i \delta(\mathbf{x}_k - \mathbf{x}_k^i).$$

The weights are normalized such that $\sum_i w_k^i = 1$. The set of particles and weights are then propagated through time to give an approximation of the filtering distribution at subsequent time steps. It only requires the definition of a suitable proposal distribution from which new particles can be simulated, and the ability to evaluate the likelihood and dynamic models. For very large number of samples this MC characterization becomes an equivalent representation to the usual functional description of the posterior pdf.

Having this kind of representation, the expectation of a function $f(.)$ can be approximated as a weighted average

$$\int f(\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{z}_{1:k}) d\mathbf{x}_k \approx \sum_{i=1}^{N_s} w_k^i f(\mathbf{x}_k^i). \tag{5}$$

For example, the mean can be estimated using $f(\mathbf{x}_k) = \mathbf{x}_k$ and the variance using $f(\mathbf{x}_k) = \mathbf{x}_k \mathbf{x}_k^T$.

2.3. Mixture Tracking

The filtering distributions in multi-object tracking applications are multi-modal. Generally, MC methods are poor at consistently maintaining the multi-modality in the filtering distribution. In practice it frequently occurs that all the particles quickly migrate to one of the modes, subsequently discarding other modes. Multiple modes arise if there is ambiguity about the object state due to insufficient measurements or if the measurements come from multiple objects. In the first case it is desirable to track all the modes until the ambiguity can be naturally resolved, and in the second, it is often required to track all the objects present.

To capture the multi-modal nature, which is inherent to our application in which tracking of multiple objects is required, the filtering distribution was modeled as an M-component mixture model

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) = \sum_{m=1}^{M} \pi_m(k) p_m(\mathbf{x}_k|\mathbf{z}_{1:k}) \tag{6}$$

with $\sum_{m=1}^{M} \pi_m(k) = 1$ and a non-parametric model is assumed for the individual mixture components. This non–parametric mixture representation can be updated in the same fashion as the two-step approach for standard Bayesian sequential estimation [6]. This approach allows for a deterministic spatial reclustering procedure $(C_k', M') = F(\mathbf{x}_k, C_k, M)$ which takes as input the particles and the current mixture representation (component indicator and number of components) and computes a new mixture representation with possibly a different number of components. Such a function encapsulates any mixture computation operation of interest, including merging, splitting, reclustering of the tracked objects.

3. EXPERIMENTAL RESULTS AND DISCUSSION

3.1. Dynamic Model

The object motion model (1) describes the evolution of the object state with respect to time. For practical implementation we adopt a nearly constant velocity model for object motion [7] and a random walk model for object intensity. In this case, the corresponding state-space model, with the state vector $\mathbf{x}_k = (x_k, \dot{x}_k, y_k, \dot{y}_k, z_k, \dot{z}_k, I_k)^T$ and zero-mean Gaussian noise ξ_k with covariance $\mathbf{Q}$, is given by

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \xi_k = \text{diag}[\mathbf{F}_1, \mathbf{F}_1, \mathbf{F}_1, 1]\mathbf{x}_k + \xi_k \quad (7)$$

where $\mathbf{Q} = \text{diag}[\mathbf{Q}_1, \mathbf{Q}_1, \mathbf{Q}_1, q_2 T]$,

$$\mathbf{F}_1 = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix}, \qquad \mathbf{Q}_1 = \begin{pmatrix} \frac{q_1}{3}T^3 & \frac{q_1}{2}T^2 \\ \frac{q_1}{2}T^2 & q_1 T \end{pmatrix}$$

where q_1 and q_2 denote the level of process noise in object motion and intensity, respectively, and T is the sampling interval. In other words, the model (7) accommodates for small accelerations in the object motion and the fluctuation in the object intensity.

3.2. Observation Model

The observation model (2) can be defined by $p(\mathbf{z}_k|\mathbf{x}_k)$, and does not necessarily have to be stationary in time. In our study $\mathbf{z}_k = \mathbf{z}_k(x, y, z)$ represents the measured image intensity. The local curvature of the image intensity distribution can be used as an additional feature in $\mathbf{z}_k$ in order to improve the performance of the tracker. The measurements produced by the objects are characterized by a combination of convex intensity distributions in all directions and a relatively high mean intensity. Noise-induced local maxima typically exhibit a random distribution of intensity changes in all directions leading to a low local curvature [8]. Assuming additive noise $\eta_k \sim N(0, \sigma_\eta^2)$ in (2) we have

$$p(\mathbf{z}_k|\mathbf{x}_k) \propto \exp\left(-\frac{\|\mathbf{z}_k - \mathbf{h}_k(\mathbf{x}_k)\|^2}{2\sigma_\eta^2}\right) \quad (8)$$

where $\mathbf{h}_k(\mathbf{x}_k)$ is the point-spread function (PSF) of the microscope modeled as

$$\mathbf{h}_k(x, y, z; \mathbf{x}_k) = b + (I_k - b)\exp\left(-\frac{(x - x_k)^2 + (y - y_k)^2}{2\sigma_{xy}^2} - \frac{(z - z_k)^2}{2\sigma_z^2}\right)$$

where b is the background intensity and σ_{xy}, σ_z represent the amount of blurring of the fluorescent tags introduced by the diffraction limited optical system. The typical approximation of the PSF by a Gaussian profile using least-squares approximation yields a value of $\sigma_{xy} \approx 80$ nm and $\sigma_z \approx 235$ nm [8]. In this case, $p(\mathbf{z}_k|\mathbf{x}_k)$ defines at time k the probability that the object with the state $\mathbf{x}_k^i$ produced a Gaussian intensity distribution in the image data.

In this study we also propose another observation model given by

$$p(\mathbf{z}_k|\mathbf{x}_k) \propto \exp\left(-\frac{(\mathbf{z}_k(x_k, y_k, z_k) - I_k)^2}{2\sigma_\eta^2}\right). \quad (9)$$

This observation model allows tracking of moving (according to (7)) intensity distributions specified in some range $I_k \in [I_{min}, I_{max}]$ avoiding hard thresholding, contrary to deterministic detection approaches. The positional information of the object in this case is calculated according to (5).

3.3. Synthetic Data

For the performance evaluation of the tracker a synthetic data generator was build. The generator creates blurred 2D or 3D images of tags with known positional information. The movement of the tags is described by (7). The PSF was modeled according to the Gibson model which is more accurate than the Gaussian approximation [9]. The generator accommodates two main types of noise that are encountered in practical situations: Poisson noise which can be used to model, for example, the effect of the quantum nature of light on the measured data, and Gaussian noise which characterizes, for example, measurement noise in the CCD detector. Figure 1 displays intensity profiles of two generated spots with different signal-to-noise ratio (SNR). The SNR is calculated as the difference in mean intensity between the object (I_o) and background (I_b), divided by the noise of the object (σ_o) [1].

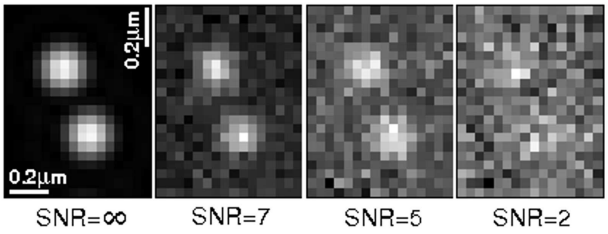

Fig. 1. Intensity profiles of generated spots for different SNR.

3.4. 2D + Time Case

An experiment was performed to determine the accuracy with which the positions of the moving tags can be estimated as a function of SNR. For that the tracking was performed on isolated spots. The statistical localization error was defined as in [8]

$$err = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\|\mathbf{x}_k^t - \mathbf{x}_k^0\|^2} \quad (10)$$

where $\mathbf{x}_k^t = (x_k, y_k, z_k)^T$ is the estimated position of the tag, which is the center position of the PSF, and $\mathbf{x}_k^0 =$

$(x_k^0, y_k^0, z_k^0)^T$ represents the true tag position known from the simulation settings. We used 10 different 2D+T datasets (size 256x256x50) with 20-50 spots per frame and ran 20 MC simulations with 10^6 particles (samples) per track. Figure 2 displays an example of synthetic data and the localization error as a function of the SNR. The localization error is in the range of $10-20$ nm for SNR$>$4 and $12-50$ nm for the range $2<$SNR<4.

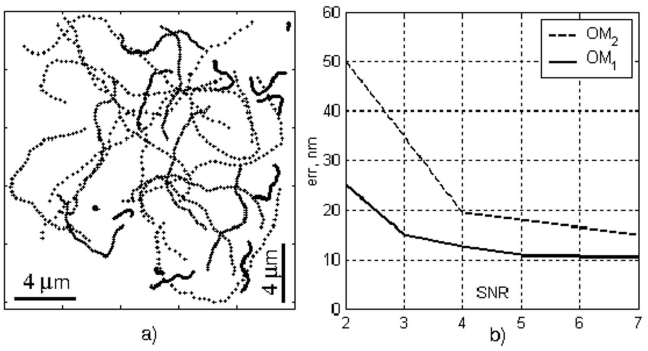

Fig. 2. (a) Synthetic data representing 20 moving tags for 50 time-frames; (b) Localization error $err(\text{SNR})$ using the models (8) denoted as OM_1, and (9) denoted as OM_2.

3.5. 3D + Time Case

The method was also tested on 3D synthetic image stacks consisting of 20 optical slices separated by 200 nm. We used 10 3D+T datasets (size 256x256x20x50) with 20-50 spots per volume and ran 20 MC simulations with 10^6 particles (samples) per track. The localization error for both observation models ranged from 60 nm (SNR=7) to 180 nm (SNR=2). The lower localization accuracy in this case was caused by approximately three times lower optical resolution of the modeled imaging system in the axial direction.

3.6. Microtubule Growth Study

The method was also evaluated on 2D image sequences of growing microtubules. Microtubules are supra-molecular structures that endow cells with characteristic morphology and motility. The ends of the microtubules are labeled with +TIPs (plus end tracking proteins) which serve as powerful markers for visualizing microtubule growth events. The images (Fig. 3) contained hundreds of tags and were taken using fluorescence confocal microscopy.

For MC simulations, the observation model (9) was used in this case, because the shape and size of the imaged microtubules (diameter$\sim$25 nm) do not allow using the observation model (8) which is applicable only for point light sources. We visually observed that the tracker successfully followed 10-20 spots originated from the manually defined regions of interest during 3-10 consequent frames until their disappearance.

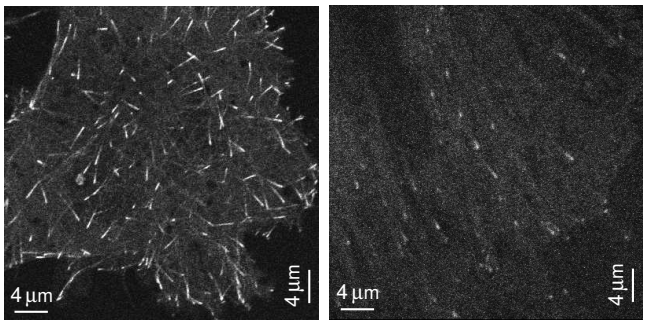

Fig. 3. Imaged microtubules labeled with plus end tracking proteins (single frames from two 2D time-lapse studies).

4. CONCLUSION

Compared to existing approaches the reported method is a substantial improvement for detection and tracking of large and time-varying numbers of spots in image data with an SNR of $2-5$. Temporal information is fully exploited, yielding more robust tracking results compared to current frame-by-frame approaches, which break down at SNR$<$5 [8]. We have demonstrated the potential of the method in the case of a real biological application. Currently we are undertaking a thorough quantitative evaluation of the method to assess its performance compared to manual tracking by expert biologists and in the case of a larger variety of biological applications.

5. REFERENCES

[1] M.K. Cheezum, W.F. Walker, and W.H. Guilford, "Quantitative comparison of algorithms for tracking single fluorescent particles," *Biophys. J.*, vol. 81, no. 4, pp. 2378–88, Oct. 2001.

[2] D. Chetverikov and J. Verestói, "Feature point tracking for incomplete trajectories," *Computing*, vol. 62, no. 4, pp. 321–38, 1999.

[3] C.B. Bergsma, G.J. Streekstra, A.W. Smeulders, and E.M. Manders, "Velocity estimation of spots in three-dimensional confocal image sequences of living cells," *Cytometry*, vol. 43, no. 4, pp. 261–72, Apr. 2001.

[4] P.Y. Burgi, A.L. Yuille, and N.M. Grzywacz, "Probabilistic motion estimation based on temporal coherence," *Neural Comput.*, vol. 12, no. 8, pp. 1839–67, Aug. 2000.

[5] M. Isard and A. Blake, "Condensation – conditional density propagation for visual tracking," *Int. J. Comp. Vis.*, vol. 29, no. 1, pp. 5–28, 1998.

[6] J. Vermaak, A. Doucet, and P. Pérez, "Maintaining multi-modality through mixture tracking," in *Proc. 9th IEEE Int. Conf. Comp. Vis.*, 2003, pp. 1110–16.

[7] X.R. Li and V.P. Jilkov, "Survey of maneuvering target tracking: Part I: Dynamic models," *IEEE T. Aero. Elec. Sys.*, vol. 39, no. 4, pp. 1333–64, Oct. 2003.

[8] D. Thomann, D.R. Rines, P.K. Sorger, and G. Danuser, "Automatic fluorescent tag detection in 3D with super-resolution: application to the analysis of chromosome movement," *J. Microsc.*, vol. 208, no. 1, pp. 49–64, 2002.

[9] S.F. Gibson and F. Lanni, "Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy," *J. Opt. Soc. Am. A.*, vol. 9, no. 1, pp. 154–66, 1992.